Status of CDF production farm upgrade

Rick Snider
for CDF production farm upgrade

Project overview
Production on the CAF
Production on the Grid

GDM meeting
August 31, 2004

Project overview

- Goals of upgrade
  - Improve resource management of production processing pipeline
    - Should allow scalable balancing of CPU and I/O resources
  - Create an extensible control system
    - Simple to add new farm management, monitoring or processing policies
  - Make production farm processes operable on other platforms
    - Short term goal: end of Fall 2004 shutdown
      - Interoperable on machines we control (CAF, dCAF)
    - Long term: staged across 2005 and 2006
      - Interoperable on machines we do not control: D0, CMS, OSG
      - Full use of common GRID tools and services with D0, CMS, OSG

Project overview

- Basic plan: phased, incremental approach
  - Bring up production at full scale on CAF
    - Use CDF/SAM analogs of Grid tools and services
    - Main point is to layer functionality in appropriate way
    - Allows access to all remote CAF resources
  - Migrate production on CAF processes to Grid
    - Full use of common GRID tools and services with D0, CMS, OSG
- Technical issues
  - Concatenation using SAM and durable storage
  - I/O using SAM + SRM + dCache
    - Load balancing between CPU and I/O
  - Gridifying
    - Job management issues
    - Pseudo-interactive services

Production on the CAF

- Two stages to production
  - Reconstruction
  - Concatenation
- Reconstruction runs on single input file
  - Configuration of executable defined by offline release
  - Output is split into multiple datasets based upon trigger content
    - Splitting configuration defined by trigger database
  - Output file sizes are generally too small for efficient storage on tape
- Concatenation
  - Concatenate intermediate files associated with particular output dataset
    - Events in a final output file define a continuous time block
      - File content specified by first run/event, last run/event
      - Allows efficient event look-up
      - Not strictly enforced (allow violations at level of a few percent)

Production on the CAF

- Currently focused on using SAM for reconstruction and concatenation
  - Based on proof of principle work at UCSD
- Reconstruction procedure currently being proto-typed
  - Query SAM to define input dataset
  - Create SAM output dataset for intermediate results
    - Reserve space in durable storage (via SRM interface)
  - Start SAM project to deliver and track files
  - Submit jobs via standard CAF submission tool
    - Pre-staging of input files or waiting for file delivery (?)
  - Write intermediate output files to durable storage
    - Use SRM interface
    - Register files in SAM

Production on the CAF

- Basic concatenation procedure under development
  - Query SAM for content, status of intermediate output
    - Reserve space for output
  - Sort file list
    - This is needed to preserve time ordering
  - Submit concatenation job
    - Write output via SAM store to dCache write pools
- Concatenation model
  - Tie concatenation job to individual input projects
    - Define relatively small, self-contained input projects
    - Concatenation resources scale as needed based upon input and output dataset
    - Requires tuning of input project definitions based upon output datasets
  - Also considering independent concatenation threads
    - Tied to individual output streams

Production on the CAF

- CPU and I/O load balancing
  - Need more flexibility in parallelizing output to tape
  - Work required to understand dCache configuration issues
    - Pool definitions, file affinities, file family definitions, etc. required to optimize throughput from farm to tape
    - Requires careful testing
- Current status
  - Reconstruction procedures are running
    - Not currently using SRM for durable storage
    - SAM interface for durable storage (sam_upload) presented by Gabrielle Garzoglio, Armando Fella, Stefano Belforte and Donatella Lucchesi under study
    - Igor Terekhov presented scheme suitable for MC production
  - Concatenation procedures and model still under development
  - dCache work not yet started

Grid enabling production jobs

- Goal
  - To run production jobs at sites we do not control
- Basic plan
  - Establish production on CAF using SAM
  - Migrate to use of SAMGrid/JIM for access to CDF resources
    - Should be “relatively easy” once production under SAM is working
      - Note that several attempts to install JIM have failed
    - Some significant technical issues to resolve
  - Enabling SAMGrid and JIM on externally controlled resources
    - This work is not CDF-specific

Grid enabling production jobs

- Technical issues
  - Job/VO management
    - CAF-specific functionality
      - Kerberos authentication
      - Condor monitoring and sandboxing features (?)
      - Assumes CDF software environment
        - *Already not needed for production processes*
    - Many CAF-specific features can be generalized using SAMGrid/JIM
  - Worker node connectivity
    - CAF provides some pseudo-interactive functionality
      - ls, head, tail, attach gdb to running executable
    - Difficult to provide this via existing gateway protocols
    - Possible solutions (??)
      - Condor glide-in
        - *Install low-level CAF environment*
      - Clarens
        - *Provides authentication, local file access via web-based interface*

Summary

- Production on CAF
  - Making progress in using SAM to run and track production, concatenation jobs
    - No significant technical hurdles expected
  - Need work to understand I/O issues using dCache
- Production on Grid
  - Migration to SAMGrid/JIM expected to be most straight-forward approach
  - Most difficult technical issue is interactive functionality
    - No clear solution has emerged

Backup slides

People

- Farm application developers
  - Suen Hou, Tsan Hsieh, Elliot Lipeles
- SAM issues
  - Krzysztof Genser, Rick St. Denis, Stephan Stonjek
- dCache issues
  - Krzysztof Genser (?)
- Project leader
  - Ashutosh Kotwal (with Ian Fisk, Rick Snider)
- Input from various people:
  - Gabriele Garzoglio, Armando Fella, Stefano Belforte, Donatella Lucchesi, Igor Terekhov, Mike Diesburg, Liz Sexton-Kennedy

Phase 1 migration plan

- Development
  - Develop new system on CAF
    - Leave existing farm intact and operating
    - Procure any dedicated hardware needed for production on CAF
      - No hardware “borrowed” from farm
  - Validate production scale capability on CAF
    - Assume dedicated services on CAF are required
- Deployment
  - Commission “official” production processing on CAF
  - Migrate and commission new system to farm
    - Requires installation of CAF software to meet short term goals
  - Move production processing to production farm
    - CAF and farm now interoperable
    - Resource sharing now possible: no more unused cycles on the farm